

Article

Identification of Cutting Workpiece Surface Defects Based on an Improved Single Shot Multibox Detector

Zhenjing Duan¹, Shushu Xi², Shuaishuai Wang¹, Ziheng Wang¹, Peng Bian¹, Changhe Li², Jinlong Song¹ and Xin Liu^{1,*}

¹ State Key Laboratory of High-Performance Precision Manufacturing, Dalian University of Technology, Dalian 116024, China; zhenjing456@126.com (Z.D.); wangshuais2022@126.com (S.W.); wzh93123456@gmail.com (Z.W.); bian_peng2023@163.com (P.B.); songjinlong@dlut.edu.cn (J.S.)

² School of Mechanical and Automotive Engineering, Qingdao University of Technology, Qingdao 266520, China; xss5284@163.com (S.X.); sy_lichanghe@163.com (C.L.)

* Corresponding author. E-mail: xinliu@dlut.edu.cn (X.L.)

Received: 15 September 2024; Accepted: 22 November 2024; Available online: 26 November 2024

ABSTRACT: In the mechanical cutting process, the surface defects of the workpiece are an important indicator of cutting quality and also reflect the condition of both the machine tool and the cutting tool. Effective detection of defects on the surface of the workpiece plays an important role in adjusting the processing conditions promptly, reducing losses, improving the utilization rate of the workpiece, and maintaining the normal operation of the equipment. To address the challenge of detecting surface defects on workpieces, an inspection method based on an improved Single Shot Multibox Detector (SSD) model is proposed. The method simplifies the detection model and reduces the computation by proposing a DH-MobileNet network instead of a VGG16 network in the SSD structure. The inverse residual structure is also used for position prediction, and null convolution is used instead of a down-sampling operation to avoid information loss. A scanning electron microscope was used to obtain the surface image of the workpiece. A dataset of workpiece surface defects was constructed and expanded, then used to train and test the model for detecting three common types of high-frequency defects: peel-off, chip adhesion, and scratches. The effect was compared with YOLO, Faster R-CNN, and the original SSD model. The detection results show that the method can detect the defects on the surface of the workpiece more accurately and quickly, which provides a new idea for defect detection in real industrial scenarios.

Keywords: Cutting machining; Workpiece surface defects; SSD; DH-MobileNet



© 2024 The authors. This is an open access article under the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

During the machining process, due to the influence of cooling and lubrication, tool structure, machine condition, cutting parameters, and materials, a variety of different topographic features will be formed on the surface of the machined workpiece, including tool texture and surface defects on the workpiece (plastic buildup, flaking, and scratches, *etc.*) [1,2]. With the rapid development of modern industry, the requirements for workpiece quality in the machinery manufacturing industry are gradually increasing. The size and type of defects on the surface of a workpiece are directly linked to the cost, performance, and service life of the machinery and equipment. Therefore, on the one hand, rational cutting parameters and auxiliary technologies (minimal quantity lubrication, cryogenics, ultrasonic vibration, cold plasma, *etc.*) are used to reduce surface defects on the workpieces during machining [3–7]; On the other hand, effective detection of defects on the surface of the workpiece plays an important role in adjusting the machining conditions promptly, reducing wear and tear, improving the utilization of the workpiece, and maintaining the normal operation of the equipment [8].

To minimize surface defects and enhance the quality of machined workpieces, it is essential to incorporate effective cooling and lubrication during the machining process [9–13]. Minimal quantity lubrication (MQL) has emerged as an environmentally friendly, resource-saving, and efficient green auxiliary processing technology [14–17]. Liu et al. combined cryogenics with MQL (biological lubricant) to grind titanium alloys, and the values of normal force,

tangential force, specific abrasive energy, coefficient of friction, temperature, and surface roughness of cryogenics-MQL were reduced by 47.6%, 44.3%, 45.1%, 19.8%, 43.9%, and 46.3%, respectively, compared with dry grinding [2]. Li et al. conducted milling experiments on titanium alloys using MQL-assisted technology, which showed a significant reduction in tool wear, probably due to the formation of an oil film on the tool and workpiece by the atomized lubricant, which significantly reduced friction and chip adhesion [18]. In order to improve the machining accuracy of difficult-to-machine materials, improve the surface quality of workpieces, and extend the service life of cutting tools, researchers have proposed the nano-lubricants minimal quantity lubrication technology (NMQL), which add nanoparticles that can play a role in cooling and lubrication to bio-oil to achieve micro-lubrication synergistic effects [19–21]. Kamal Kishore et al. conducted experiments on cutting Inconel 625 employing NMQL-assisted technology, which showed a significant reduction in surface defects on the workpiece [22].

Adopting certain methods to reduce workpiece surface defects during the cutting process has been well studied. However, detecting surface defects on the cut workpiece is equally important. In recent years, defect monitoring technologies based on deep learning and machine vision have been increasingly adopted in industrial production, significantly advancing the development of the intelligent manufacturing industry [23,24]. Deep learning algorithms have demonstrated state-of-the-art properties in classification and target detection tasks, and their use in surface defect detection is also widespread [25,26]. Deep learning, as a simple and fast method of data analysis, has made breakthroughs in areas such as image processing and target detection [27]. Xie et al. proposed deep learning based on a Convolutional Neural Network (CNN) structure for surface defect detection of magnetic tiles. In their investigation, AlexNet, Visual Geometry Group 16 (VGG16), and ResNet50 were utilized as feature extraction models. The method utilizes a traditional CNN architecture with multiple convolutional layers followed by a Softmax layer to make the final decision [28]. Tabernik et al. proposed a two-stage approach for surface anomaly segmentation based on deep learning models. The researchers developed a segmentation network that consists of 11 convolutional layers trained on pixel labels. Notably, their proposed model shows efficient learning even with a limited number of defective surfaces, as it achieves satisfactory results with only about 25–30 defective training samples [29]. Neuhauser et al. proposed a CNN-based model for classifying and detecting surface defects on extruded aluminum profiles. The model uses Faster R-CNN and ResNet50 as the feature extraction backbone to capture relevant defect features. The study demonstrated the successful application of the proposed method in accurately classifying and detecting surface defects on extruded aluminum profiles [30]. Le et al. proposed an end-to-end lightweight deep learning model for surface defect detection of industrially manufactured parts anchored by YOLOv5. They employed coordinated attention in the Cross Stage Partial (CSP) module to enhance feature extraction [31]. Niu et al. proposed a novel data enhancement strategy to enhance the focus on low-confidence regions in downstream CNNs. The CNN-based segmentation deep learning model was further improved by employing the newly proposed data enhancement method utilizing confidence maps. Experiments demonstrated the model's high performance in accurately segmenting defects with weak features [32]. Yu et al. introduced a defect detection model for metal surface defects using a small sample learning approach. Their method utilized the matrix decomposition attention mechanism to solve the problem of limited defect samples in industrial production lines. This method could effectively detect and classify metal surface defects using a small amount of labeled data [33].

To detect workpiece surface defects more accurately and efficiently, Scanning Electron Microscope (SEM) images of the milled workpiece surfaces, captured under both dry and NMQL-assisted conditions, were used as training sample data. Under dry conditions, the workpiece surface exhibited more and larger morphological defects, whereas the number of surface defects was fewer under NMQL-assisted conditions. The sample data from these two working conditions gave variability in the number and size of defects on the surface of the workpiece, which helped the model to adapt to the identification of workpiece surfaces with different qualities. In this paper, we propose a workpiece surface defect detection method based on an improved Single Shot MultiBox Detector (SSD) network model, which proposes a DH-MobileNet network based on the MobileNet network and replaces the VGG16 in the SSD model with this network. To address the problem of information loss caused by transmission between subsequent convolutional layers of the SSD model, the inverse residual structure is introduced, while to improve the accuracy of detection of small targets, the null convolution is used instead of the downsampling operation in the inverse residual structure. SEM was utilized to measure high-resolution images to obtain the characteristics of defects on the surface of the workpiece. By optimizing the dataset and training the network model to effectively detect surface defects on workpieces, this approach offers a new solution to the challenge of defect detection in industrial production.

2. SEM Sample Data

Due to the small number of publicly available datasets on workpiece defects, a self-tested dataset was used for this experiment. The sample dataset was tested using SEM. Due to the large amount of sample data, the sample data used includes SEM images from previous tests (Figure 1). A camera can also be used to collect data in practical applications. The test workpieces were processed under dry and NMQL conditions, respectively. The sample material used in this study is primarily aluminum alloy 7050, with its composition and physical properties detailed in Tables 1 and 2.

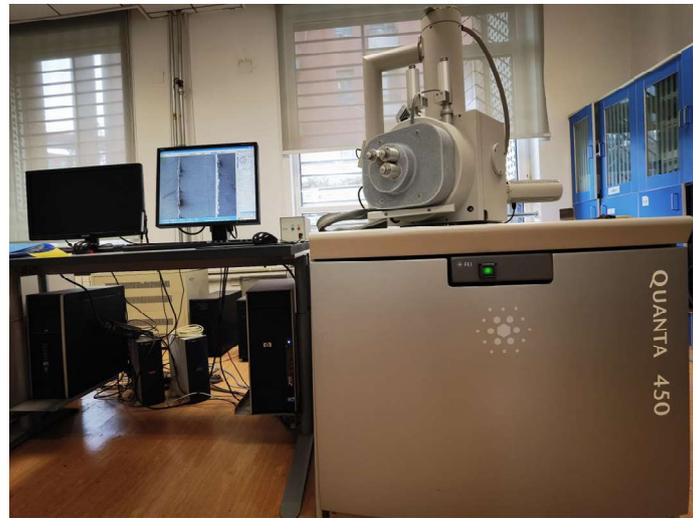


Figure 1. Sample data collection tools.

Table 1. Chemical composition of workpiece.

Element	Al	Cr	Zr	Zn	Si	Fe	Mn	Mg	Ti	Cu
Component/%	margin	0.04	0.1	5.7	0.12	0.1	0.10	2.1	0.12	2.2

Table 2. Mechanical properties.

Tensile Strength (MPa)	Yield Strength (MPa)	Hardness (HB)	Elongaon (%)	Density (g/m ³)
552	489	140	11	2.83

3. SSD Target Detection Model

3.1. Model Structure

SSD is a target detection algorithm proposed by Wei Liu in ECCV 2016, which utilizes different convolutional layers for boundary and classification regression to achieve better detection and speed [34–36]. The SSD network structure is divided into two parts. The first part is the VGG16 network structure with high classification accuracy and its classification layer removed. The second part is a network structure constructed by using convolutional layers instead of two fully connected layers and adding four more convolutional layers to incorporate feature pyramid detection for multi-scale target detection [37,38]. The network model structure of SSD is shown in Figure 2.

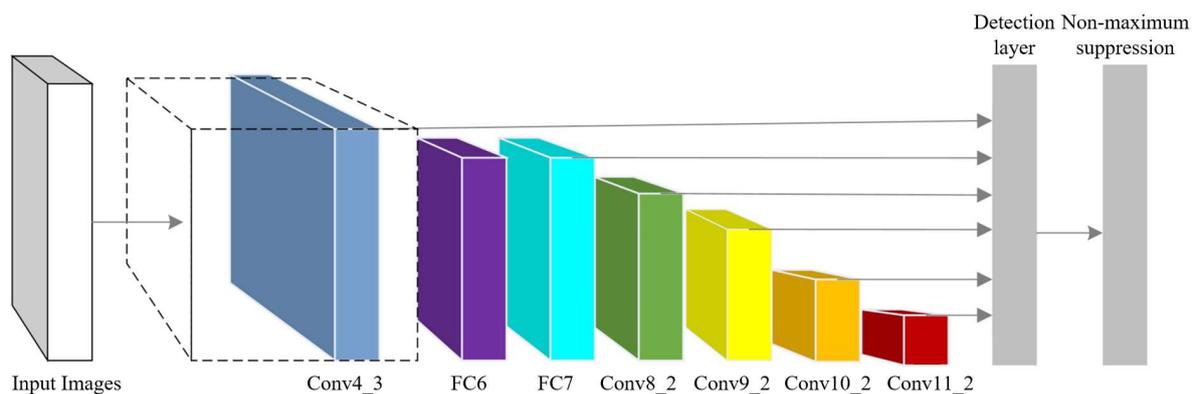


Figure 2. Structure of SSD network model.

The SSD consists of a backbone network and a multi-scale feature layer, with the backbone network (VGG-16) used for feature extraction. The model detects objects at multiple scales to capture objects of varying sizes. Where Conv4_3 has a higher resolution and is suitable for detecting small objects; FC6 is one of the fully connected layers of the VGG-16 network and is located in the deeper layers of the network; FC was changed to a convolutional layer to preserve the spatial resolution Conv8_2, Conv9_2, Conv10_2, and Conv11_2 are additional convolutional layers that were added in the backbone network (e.g., VGG-16) after it, which are used to construct multi-scale feature maps. These feature layers sequentially have smaller and smaller spatial resolutions.

3.2. Scale Calculation

In the modeling framework of SSD, each selected feature map has k frames with different size and width ratios, which are called default frames. k denotes the number of different frames at each position. The default frames on the feature maps of different solution layers are shown in Figure 3.

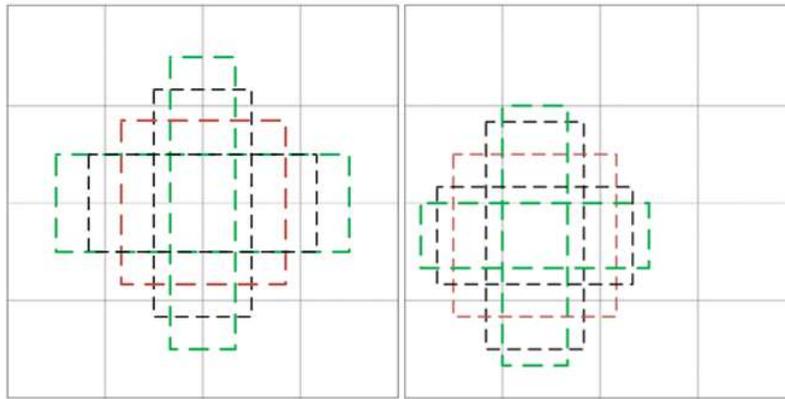


Figure 3. Characteristic Graphs of Outputs of Different Convolutional Layers.

Each default box predicts the number of targets as B and four positional parameters. In this experiment, B is set to 3 because there are three typical defects in the surface of the workpiece, and the calculation of the default box for each feature map is shown in Equation (1).

$$P_i = P_{\min} + \frac{P_{\max} - P_{\min}}{n - 1} (i - 1), i = 1, 2, \dots, n \quad (1)$$

where n is the number of feature maps, P_{\min} and P_{\max} are parameters that can be set, which are taken as 0.2 and 0.9 respectively in this experiment, meaning that the scale of the bottom layer is 0.2 and the scale of the top layer is 0.9, and P_i denotes the scale of the i th feature map. To control the fairness of the feature vectors in the training and testing experiments, the same five aspect ratios $b = (1, 2, 3, 0.5, 0.33)$ were used to generate the default boxes, each of which can be described as:

$$W_i^b = P_i \sqrt{b} \quad (2)$$

$$H_i^b = \frac{P_i}{\sqrt{b}} \quad (3)$$

Equation (2) represents the width of the default box, and Equation (3) represents the height of the default box. When the aspect ratio is 1, it is necessary to add a default box scaled to $P_i' = \sqrt{P_i P_{i+1}}$. Each default box is centered on $(\frac{a+0.5}{|g_x|}, \frac{a+0.5}{|g_x|})$, and $|g_x|$ is the size of the i th feature map, $a, b \in \{0, 1, \dots, |g_x|\}$.

3.3. Loss Function

The SSD model is trained for regression on both target location and species, and the training process generates a loss volume, which is calculated using a loss function. The overall objective loss function of the SSD model is the weighted sum of the location loss function (loc) and the confidence loss function (conf), As shown in Equation (4) [39]:

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + aL_{loc}(x, l, g)) \tag{4}$$

where the parameter a is used to balance the confidence loss and position loss; c is the classification confidence; l is the offset of the prediction frame, including the translation offset of the center coordinates and the scaling offset of the width; g is the calibration frame for the actual position of the target; and N is the number of default frames matching the calibration frames of the category. When N equals 0, it means that the loss is 0.

The position loss function is the $smooth_{L1}Loss$ between the parameters of the prediction box l and the true labeled value box g . The regression to the bias of the default border d centered at (cx, cy) and its width (w) height (h) is defined as shown in Equation (5):

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k smooth_{L1}(l_i^m - g_j^m) \tag{5}$$

where $smooth_{L1}(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & otherwise \end{cases}$, l_i^m is the prediction box, and g_j^m is the true box.

The confidence loss function is defined as shown in Equation (6).

$$L_{locf}(x, c) = \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0) \tag{6}$$

where \hat{c}_i^p is $\frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}$, which is generated through Softmax, which is the activation function.

Unlike the original MultiBox matching method, the matching strategy during SSD model training is to first match the default bounding box with the real bounding box. If the Intersection over Union (IoU) of the two is higher than the threshold (the default threshold is 0.5), then the default detection box matches with the real box; on the contrary, if the overlap rate is less than 0.5, then the two are not matched match [40].

4. Design and Optimization of Experimental Models

The SSD model shows better performance in target detection, taking advantage of multiple convolutional layers for target detection applications. However, the SSD model requires more parameters in the detection process as well as the subsequent convolutional layers ignore the connection relationship between layers, thus increasing the computational volume, while the detection of small targets is poor and is prone to problems such as leakage and false detection. The MobileNet network uses fewer parameters, which reduces training time. However, this comes at the cost of lower detection accuracy. To improve the network accuracy while reducing the training parameters, firstly, the DH-MobileNet network is proposed by combining Dilated Convolution and Hierarchical Feature Fusion and used to replace VGG16 in the SSD model; then, the inverse residual structure is utilized to improve the subsequent convolutional layers of the SSD model by using the inverse residual structure to improve the subsequent convolutional layers of the SSD model. The structure of the optimized SSD model is shown in Figure 4.

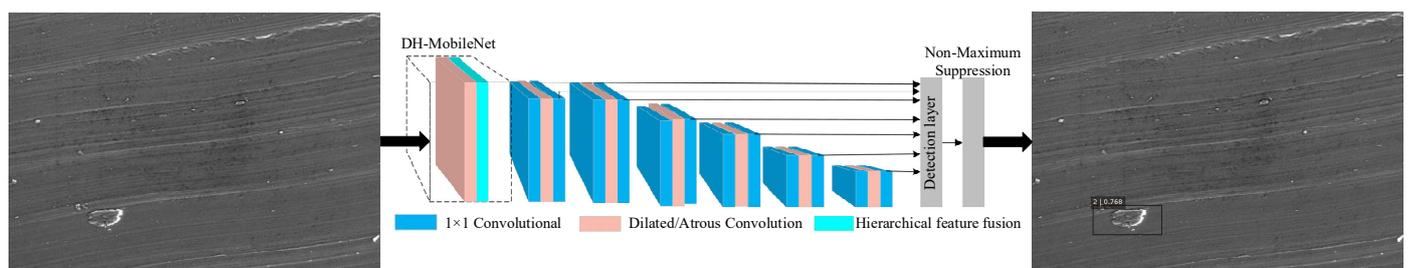


Figure 4. Optimized SSD Model.

The optimized model could be improved by adding a dilation layer to the convolutional network to enhance the model’s receptive field. Dilated Convolution extended the range of the receptive field by inserting intervals between

the elements of the convolutional kernel without increasing the number of parameters or decreasing the resolution of the feature map.

4.1. DH-MobileNet Network Feature Extraction Principle

MobileNet was a lightweight convolutional neural network proposed by Howard et al. The basic unit is a deeply separable convolution, which requires fewer parameters for training and can save and enable the deployment of mobile devices [41]. The structure of the MobileNet model is shown in Figure 5.

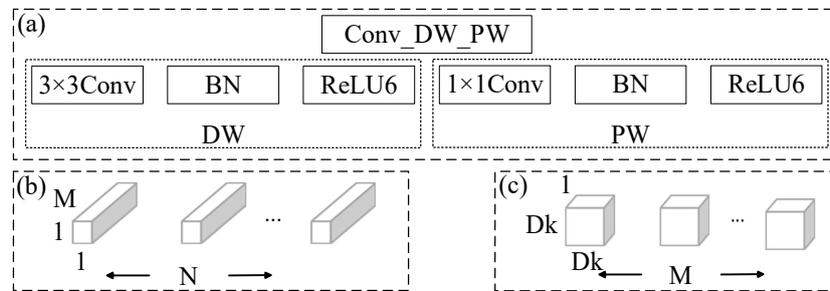


Figure 5. MobileNet structure (a) Network structure of MobileNet; (b) Depth-separable convolutional structure; (c) Point convolution kernel.

Figure 5a shows the network structure of MobileNet, where Conv-Dw-Pw is a depth-separable convolutional structure, which consists of two parts: depth-wise (DW) and point-wise (PW) convolution. The DW is constructed using a 3×3 deep convolutional kernel, as shown in Figure 5b; The PW is composed using a 1×1 point convolution kernel, as shown in Figure 5c. BN stands for batch normalization, and ReLU6 for activation function.

During the model training process, the receptive field is expanded through pooling operations to enhance detection accuracy. Still, the pooling operation will cause the problem of partially missing image information. Dilated Convolution can increase the receptive field of the input image without losing the image information; therefore, introducing the Dilated Convolution (DC) operation into the MobileNet network not only ensures the integrity of the image information but also increases the range of the receptive field of the convolutional layer, thus ensuring the detection accuracy [42]. A cavity convolution operation with a convolution kernel size of 3×3 and a dilation rate of 2 is shown in Figure 6, where nine points in a 7×7 region are convolved with a 3×3 convolution kernel, the rest of the points contain a parameter set to 0, and the size of the sensory field after convolution is $F_{dilation} = \left[2^{\left(\frac{dilation}{2}\right)+2} - 1 \right]$.

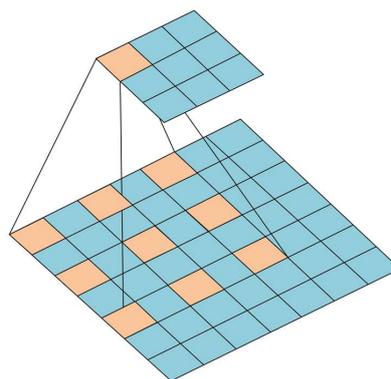


Figure 6. Dilated Convolution operation.

Hierarchical Feature Fusion (HFF) is to sum the outputs of each convolutional unit of a null convolutional layer in turn and apply the concatenation operation to each summed result to obtain the final output, which can expand the learning parameters without increasing the complexity of the network and strengthen the network continuity [40]. The HFF structure is shown in Figure 7.

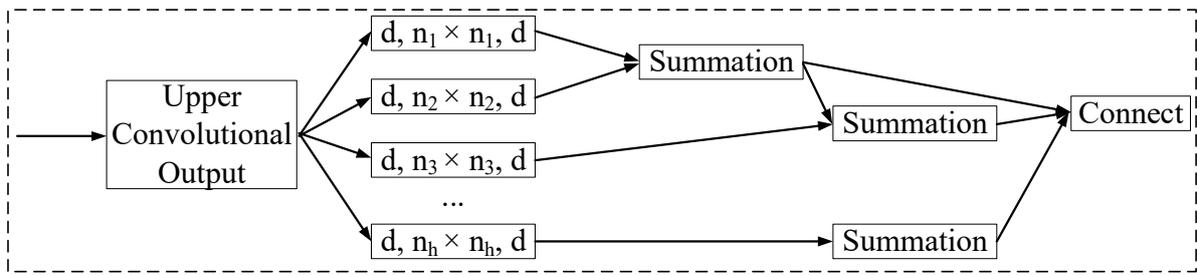


Figure 7. Hierarchical Feature Fusion.

The depth-separable convolutional structure of DH-MobileNet requires only 1/34 of the VGG16 network parameters to achieve the same classification accuracy in the classification task, and the same output as the standard convolution can be obtained through operations such as dimensionality reduction, which can reduce the amount of computation and improve the effect of training. The network structure of DH-MobileNet is shown in Figure 8.

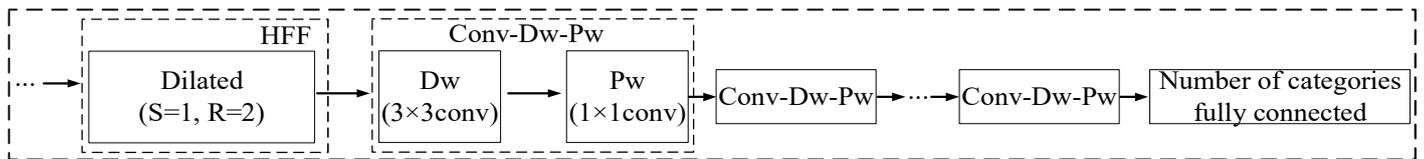


Figure 8. Architecture of DH-MobileNet.

In Figure 8, Dilated denotes the cavity convolution, R denotes the dilation rate, and Conv-Dw-Pw represents the depth-separable convolution structure.

4.2. SSD Convolutional Network Layer Optimization

Since the subsequent convolutional layers in the SSD model ignore the connectivity between layers, this leads to poor detection of small targets. Therefore, the introduction of the inverse residual structure reduces the loss of information caused by nonlinear transformations at low latitudes during the learning process. To prevent the loss of feature information during the down-sampling process, a depthwise separable convolution is used to replace traditional down-sampling in the inverse residual structure. Additionally, the ReLU activation function is replaced with the higher-performance ReLU6. The ReLU6 activation function can be expressed as: $Y = \min(\max(X, 0), 6)$. Where Y is the output of ReLU6 and X is the value of each pixel in the feature map. Batch normalization (BN) algorithm is used for normalization. The BN algorithm supports automatic data distribution adjustment and prevents gradient vanishing and complex variable adjustment by setting two learning parameters [43]. The inverse residual structure after adding the null convolution is shown in Figure 9.

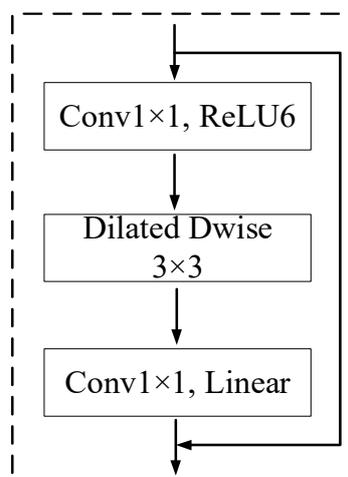


Figure 9. Reverse residual structure before and after improvement.

5. Experimental Design

The configuration of the computer used for the experiments is as follows: the operating system is Windows 10, the CPU model is Intel(R) Core(TM) i7-8750H, the frequency of the CPU is 2.20 GHz, the RAM is 8.0 GB, and the model of the graphic card (GPU) is GeForce GTX 1060, and TensorFlow is set up based on the above hardware environment on the above hardware.

The process of training the model starts with training the training set and eliminates the discarded below-threshold detection frames using a non-great-value suppression algorithm [43]. After iterating and adjusting the parameter settings to optimize the model’s performance, it is subsequently tested for accuracy using the test set from the experimental dataset. The model’s performance is examined by analyzing the detection accuracy, detection speed, and number of parameters of the model in the test set.

6. Results and Discussion

6.1. Structure of Data Samples

Three common defects (chip adhesion, flaking, and pear gouge) were selected for this experiment. In this paper, while satisfying the number of samples and fully considering the specification factor of the samples, images containing three kinds of defects: chip adhesion, flaking, and pear gouge are collected by SEM, such as typical defects on the surface of the workpiece as shown in Figure 10.

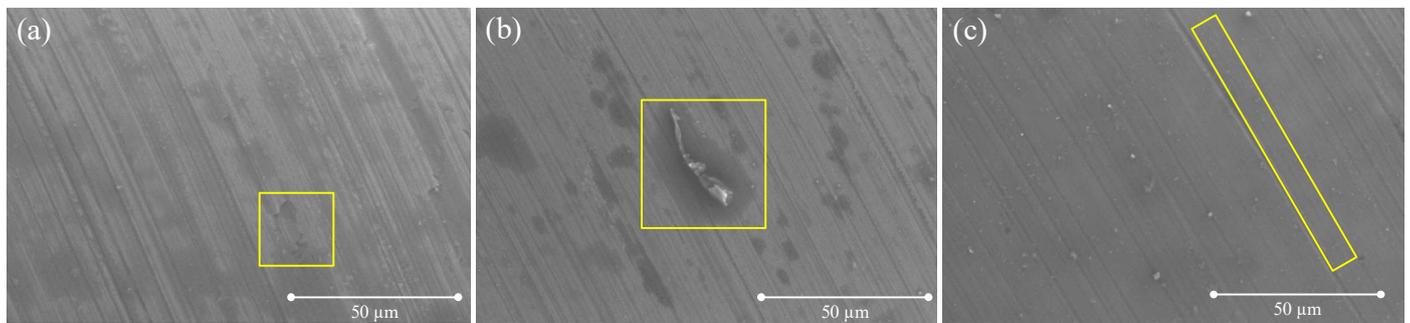


Figure 10. Surface defects in cutting processing (a) Peel off; (b) Chip adhesion; (c) Scratch.

The collected defective images are subjected to data expansion, and a data set is established. In data expansion, the main operations performed are rotation, horizontal migration, vertical migration, and scaling in four ways, and the specific implementation method of data expansion is shown in Table 3.

Table 3. Data expansion implementation methods.

Data Extension Methods	Change
Rotation	Rotate the defective image from 0° to 10°
Horizontal migration	Horizontally offset the image by 10% of the image length.
Vertical migration	Offset the image vertically by 10% of the image length.
Scaling	Enlarge the defective image to a 10% scale.

The number of training samples is increased after data augmentation, which strongly improves the training performance of the dataset. 80% of the data-augmented dataset is used as a training set and 20% as a test set.

6.2. Model Accuracy Analysis

6.2.1. Convergence Rate Analysis

The proposed DH-MobileNet-SSD model was compared with three models, Faster R-CNN, SSD, and YOLO, in the same dataset for experiments. Each experimental model was iterated 36,000 times, as this number was deemed sufficient based on hardware limitations and practical application requirements. Beyond 36,000 iterations, further training showed minimal significance since the loss values had already stabilized. The loss function image during training is shown in Figure 11.

As can be seen from Figure 11, during the training of the dataset of this experiment, the loss values of each model converged continuously with the increase of the number of iterations. After 25,000 iterations, each model stabilized. Among them, the DH-MobileNet-SSD could better accelerate the convergence speed of the network.

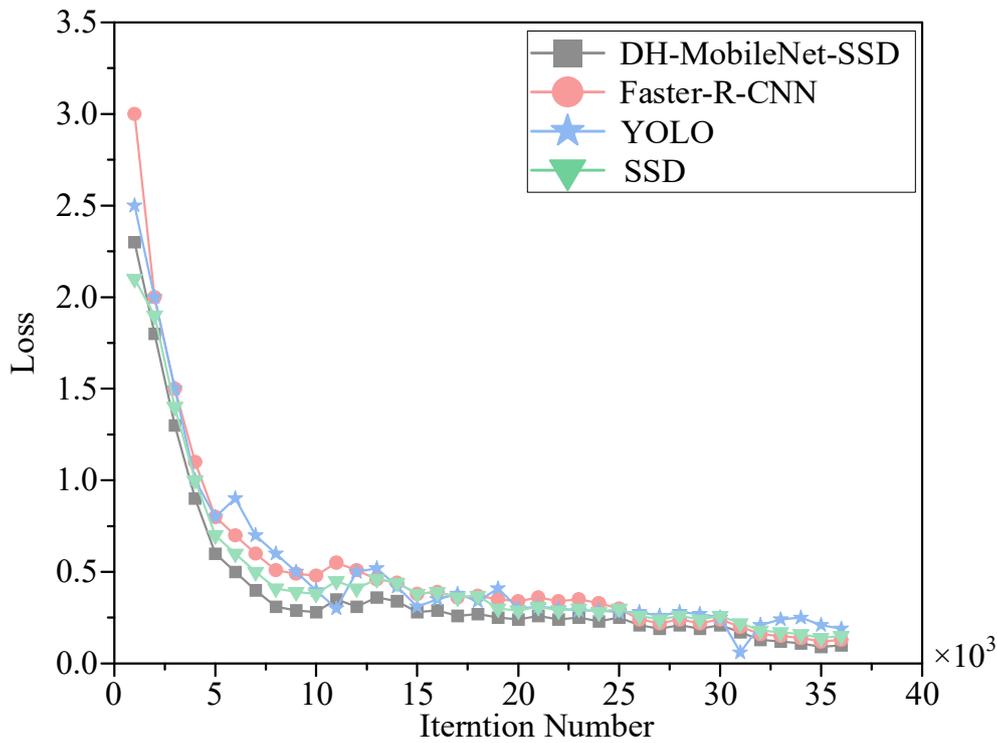


Figure 11. Loss function training image.

6.2.2. Performance Evaluation

Precision, Average Precision (AP) and Mean Average Precision (mAP) were the main commonly used parameters to measure the performance of a target detection model, where precision was the proportion of defect types that were truly judged to be accurate in the images judged to be accurate; Average Precision was the average of the precision, reflecting the average detection accuracy of each defect; mean Average Precision represents the average accuracy of the three categories of defects, reflecting the effectiveness of the model in detecting the entire dataset. The formulas for the three commonly used parameters are shown in Equations (7)–(9).

$$\text{Precision} = \frac{\text{Num}(\text{TruePosition})}{\text{Num}(\text{TotalObjects})} \tag{7}$$

$$\text{AP} = \frac{\sum \text{Precision}}{\text{Num}(\text{TotalImages})} \tag{8}$$

$$\text{mAP} = \frac{\sum \text{AP}}{\text{Num}(\text{Classes})} \tag{9}$$

The test set was detected using the trained network. The test set images contain three different types of defect samples, each with 60 different images. The results of the precision of the trained detection network for the three defective samples are shown in Table 4.

Table 4. Precision table for detection of three types of defects.

Defect Types	Sample Size	Number of Successful Tests	Number of Omissions	Number of Error Detections	Precision %
Chip adhesion	60	56	4	0	93.3
Peel off	60	51	4	5	85
Scratch	60	53	3	4	88.3
Total	180	160	11	9	88.9

The inspection results showed that among the three types of defects, chip adhesion was detected with higher accuracy, but it also had a higher number of missed detections. Both types of defects, spalling and pear gouge, are still missed and wrongly detected. Overall, it can be concluded that the DH-MobileNet-SSD model has a higher detection accuracy of 88.9% for these three defects.

The Average Precision (AP) values of the DH-MobileNet-SSD model with the three models Faster-RCNN, SSD, and YOLO at the same number of iteration steps are shown in Figure 12.

Figure 12 shows that the AP value of the DH-MobileNet-SSD model is much higher than that of the Faster R-CNN, SSD, and YOLO models, which further validates that the model used has better results in the detection of defects on the surface of workpieces.

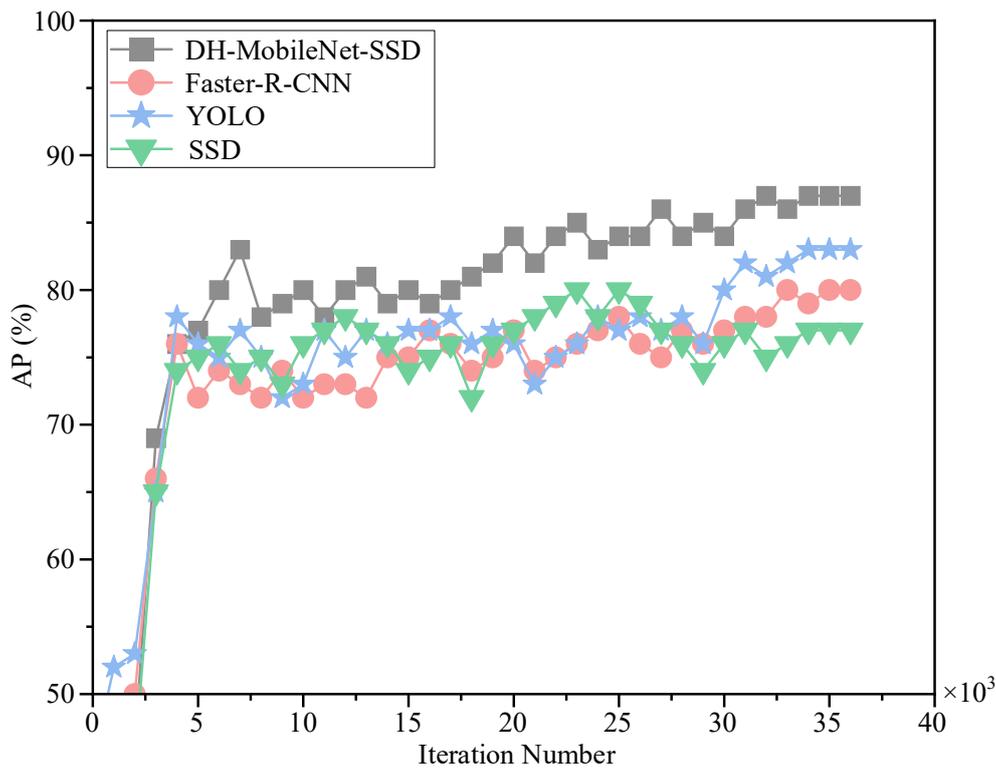


Figure 12. Detection results of defect AP on the workpiece surface.

With the same number of training steps and the same training set, the DH-MobileNet-SSD model used in this paper is compared with the Faster R-CNN, SSD, and YOLO models in terms of mAP and average detection time, and the results are shown in Figure 13.

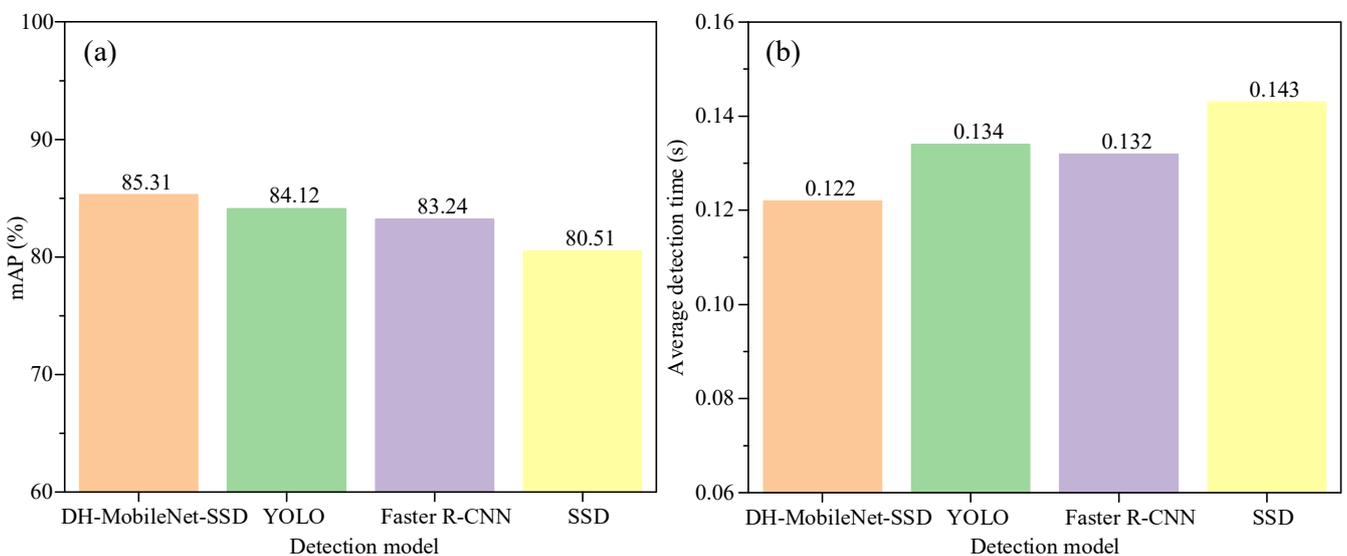


Figure 13. Comparison of defect detection methods (a) mAP; (b) Average detection time.

From Figure 13, it can be seen that the average accuracy mean of DH-MobileNet-SSD is larger and more accurate than the average accuracy mean of the Faster R-CNN-based model, YOLO-based model, and SSD-based model. Analyzing the detection speed, the average detection time of this paper's model is 0.122 s, which is lower than the other three detection algorithms, indicating that this paper's method is more time-saving than the Faster R-CNN model, SSD model, and YOLO model.

6.3. Data Set Sample Detection

One type of defect was selected to evaluate detection performance across the dataset using four models. The detection results are presented in Figure 14. Specifically, Figure 14a displays the detection results of the Faster R-CNN model, Figure 14b shows the results of the YOLO model, Figure 14c presents the detection results of the original SSD model, and Figure 14d illustrates the detection results of the experimental model, DH-MobileNet-SSD. The detection plots reveal that the improved SSD model (DH-MobileNet-SSD) achieves higher detection accuracy, with the detection frames aligning more precisely with the defective regions compared to the other models.

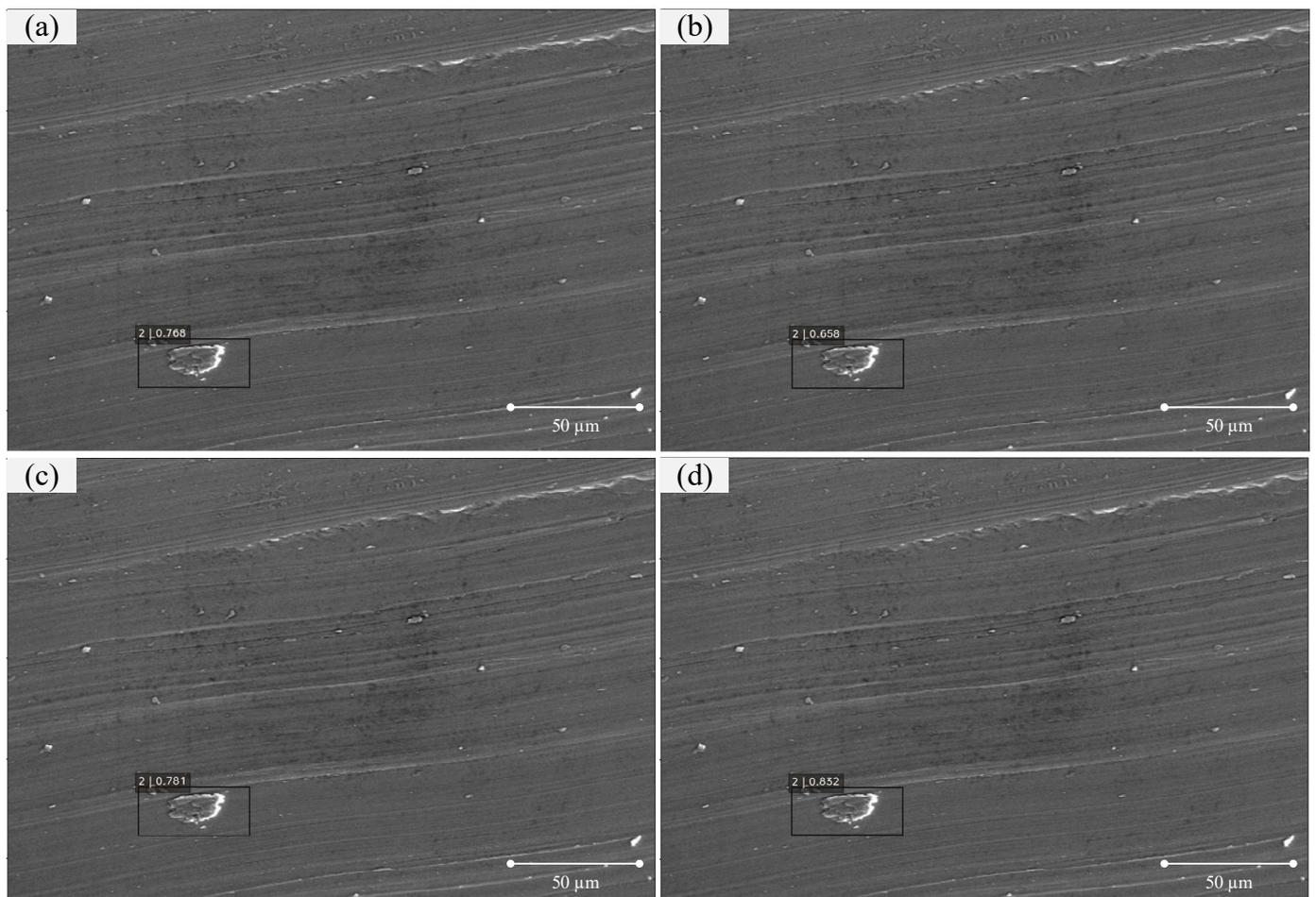


Figure 14. Detection Result Diagram (a) SSD; (b) YOLO; (c) Faster R-CNN;(d) DH-MobileNet-SSD.

The three defects were detected separately using the DH-MobileNet-SSD model, and the detection results are shown in Figure 15, with Peel off in Figure 15a, Chip adhesion in Figure 15b, and Scratch in Figure 15c. It could be seen from the detection plots that three defects can be detected with a high rate of correctness. Therefore, the DH-MobileNet-SSD model proposed in this study demonstrates superior detection performance and exhibits good stability when applied to this dataset.

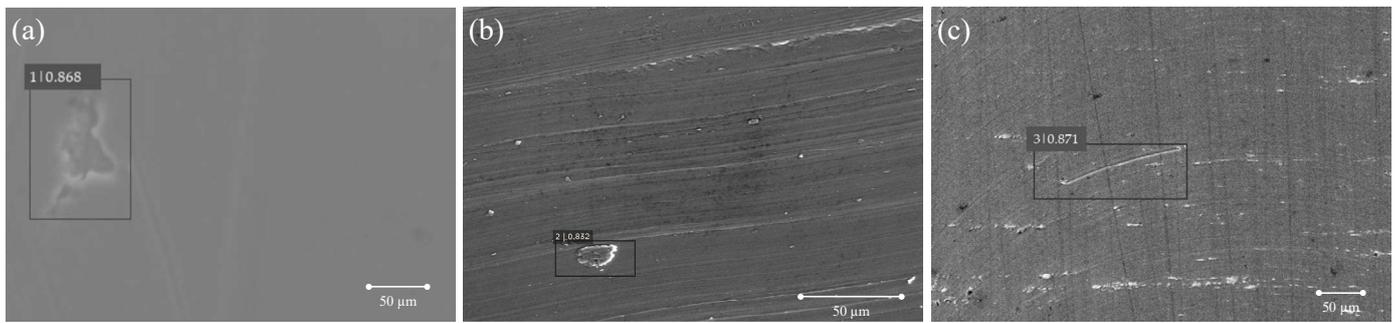


Figure 15. Three Defect Detection Results (a) Peel off (b); Chip adhesion; (c) Scratch.

7. Conclusions

To improve the detection of defects on the surface of the cut workpiece, Scanning Electron Microscope (SEM) images of the mechanical workpiece surface were first acquired. Then DH-MobileNet-SSD was proposed as a method to detect the defects on the surface of the workpiece based on the SSD network model for three kinds of high-frequency defects (peel off, chip adhesion, and scratch). The method utilized the deep convolutional features, effectively avoiding the problem of traditional target detection models relying on manual features. The conclusion is as follows:

- (1) The DH-MobileNet-SSD inspection model proposed in this paper can effectively detect surface defects (peel off, chip adhesion, and scratch) on workpieces.
- (2) DH-MobileNet-SSD achieves 88.9% precision in workpiece defect detection, which is higher than Faster R-CNN, YOLO, and the original SSD method under the same conditions. However, the training network is not yet fully optimized; its accuracy requires further improvement, and its overall performance still needs enhancement. This remains an area for future research and development.
- (3) The mAP of the DH-MobileNet-SSD model (85.31%) is higher and more accurate compared to the mAP of the Faster R-CNN-based model, the YOLO-based model, and the SSD-based model. Analyzed in terms of detection speed, the average detection time of DH-MobileNet-SSD is 0.122 s, which is lower than the other three detection algorithms.

Author Contributions

Z.D.: Preparation, Experiment, Writing. S.X.: Programming, Writing. S.W.: Visualization. Z.W.: Visualization. P.B.: Project administration. C.L.: Methodology. J.S.: Formal analysis. X.L.: Funding acquisition, Supervision, Resources.

Ethics Statement

Not applicable.

Informed Consent Statement

Not applicable.

Funding

This research was funded by [Fundamental Research Funds for the Central Universities] grant number [DUT19ZD202] and [National Natural Science Foundation of China] grant number [52475430].

Declaration of Competing Interest

The authors declare that they have no known competing interests or personal relationships that may affect the work reported in this paper.

References

1. Yang M, Hao JC, Wu WT, Li ZH, Ma YQ, Zhou ZM, et al. Critical cutting thickness model considering subsurface damage of zirconia grinding and friction—Wear performance evaluation applied in simulated oral environment. *Tribol. Int.* **2024**, *198*, 109881. doi:10.1016/j.triboint.2024.109881.
2. Liu MZ, Li CH, Yang M, Gao T, Wang XM, Cui X, et al. Mechanism and enhanced grindability of cryogenic air combined with biolubricant grinding titanium alloy. *Tribol. Int.* **2023**, *187*, 108704. doi:10.1016/j.triboint.2023.108704.

3. Duan ZJ, Li CH, Zhang YB, Dong L, Bai XF, Yang M, et al. Milling surface roughness for 7050 aluminum alloy cavity influenced by nozzle position of nanofluid minimum quantity lubrication. *Chin. J. Aeronaut.* **2021**, *34*, 33–53. doi:10.1016/j.cja.2020.04.029.
4. Cao Y, Ding WF, Zhao BA, Wen XB, Li SP, Wang JZ. Effect of intermittent cutting behavior on the ultrasonic vibration-assisted grinding performance of Inconel718 nickel-based superalloy. *Precis. Eng.* **2022**, *78*, 248–260. doi:10.1016/j.precisioneng.2022.08.006.
5. Cao Y, Yin JF, Ding WF, Xu JH. Alumina abrasive wheel wear in ultrasonic vibration-assisted creep-feed grinding of Inconel 718 nickel-based superalloy. *J. Mater. Process Tech.* **2021**, *297*, 117241. doi:10.1016/j.jmatprotec.2021.117241.
6. Qu SS, Yao P, Gong YD, Chu DK, Yang YY, Li CW, et al. Environmentally friendly grinding of C/SiCs using carbon nanofluid minimum quantity lubrication technology. *J. Clean. Prod.* **2022**, *366*, 132898. doi:10.1016/j.jclepro.2022.132898.
7. Liu MZ, Li CH, Zhang YB, Yang M, Gao T, Cui X, et al. Analysis of grain tribology and improved grinding temperature model based on discrete heat source. *Tribol. Int.* **2023**, *180*, 108196. doi:10.1016/j.triboint.2022.108196.
8. Wang HJ, Mo H, Lu SL, Zhao XM. Electrolytic capacitor surface defect detection based on deep convolution neural network. *J. King Saud Univ.-Comput. Inf. Sci.* **2024**, *36*, 101935. doi:10.1016/j.jksuci.2024.101935.
9. He T, Liu NC, Xia HZ, Wu L, Zhang Y, Li DG, et al. Progress and trend of minimum quantity lubrication (MQL): A comprehensive review. *J. Clean. Prod.* **2023**, *386*, 135809. doi:10.1016/j.jclepro.2022.135809.
10. Sarikaya M, Gupta MK, Tomaz I, Danish M, Mia M, Rubaiee S, et al. Cooling techniques to improve the machinability and sustainability of light-weight alloys: A state-of-the-art review. *J. Manuf. Process.* **2021**, *62*, 179–201. doi:10.1016/j.jmapro.2020.12.013.
11. Duan ZJ, Wang SS, Wang ZH, Li CH, Li YH, Song JL, et al. Tool wear mechanisms in cold plasma and nano-lubricant multi-energy field coupled micro-milling of Al-Li alloy. *Tribol. Int.* **2024**, *192*, 109337. doi:10.1016/j.triboint.2024.109337.
12. Duan ZJ, Wang SS, Li CH, Wang ZH, Bian P, Sun J, et al. Cold plasma and different nano-lubricants multi-energy field coupling-assisted micro-milling of Al-Li alloy 2195-T8 and flow rate optimization. *J Manuf Process.* **2024**, *127*, 218–237. doi:10.1016/j.jmapro.2024.07.146.
13. Zhang Y, Li L, Cui X, An Q, Xu P, Wang W, et al. Lubricant activity enhanced technologies for sustainable machining: Mechanisms and processability. *Chin. J. Aeronaut.* **2024**, doi:10.1016/j.cja.2024.08.034.
14. Jia DZ, Zhang YB, Li CH, Yang M, Gao T, Said Z, et al. Lubrication-enhanced mechanisms of titanium alloy grinding using lecithin biolubricant. *Tribol. Int.* **2022**, *169*, 107461. doi:10.1016/j.triboint.2022.107461.
15. Jia DZ, Li CH, Liu JH, Zhang YB, Yang M, Gao T, et al. Prediction model of volume average diameter and analysis of atomization characteristics in electrostatic atomization minimum quantity lubrication. *Friction* **2023**, *11*, 2107–2131. doi:10.1007/s40544-022-0734-2.
16. Cönger DB, Yapan YF, Emiroglu U, Uysal A, Altan E. Influence of singular and dual MQL nozzles on sustainable milling of Al6061-T651 in different machining environments. *J. Manuf. Process.* **2024**, *109*, 524–536. doi:10.1016/j.jmapro.2023.12.043.
17. Krishnan GP, Raj DS. Machinability and tribological analysis of used cooking oil for MQL applications in drilling AISI 304 using a low-cost pneumatic operated MQL system. *J. Manuf. Process.* **2023**, *104*, 348–371. doi:10.1016/j.jmapro.2023.09.028.
18. Li JH, Shi WT, Lin YX, Li J, Liu S, Liu B. Comparative study on MQL milling and hole making processes for laser beam powder bed fusion (L-PBF) of Ti-6Al-4V titanium alloy. *J. Manuf. Process.* **2023**, *94*, 20–34. doi:10.1016/j.jmapro.2023.03.055.
19. Cui X, Li CH, Zhang YB, Said Z, Debnath S, Sharma S, et al. Grindability of titanium alloy using cryogenic nanolubricant minimum quantity lubrication. *J. Manuf. Process.* **2022**, *80*, 273–286. doi:10.1016/j.jmapro.2022.06.003.
20. Khanna N, Agrawal C, Pimenov DY, Singla AK, Machado AR, da Silva LRR, et al. Review on design and development of cryogenic machining setups for heat resistant alloys and composites. *J. Manuf. Process.* **2021**, *68*, 398–422. doi:10.1016/j.jmapro.2021.05.053.
21. Wang XM, Li CH, Zhang YB, Said Z, Debnath S, Sharma S, et al. Influence of texture shape and arrangement on nanofluid minimum quantity lubrication turning. *Int. J. Adv. Manuf. Tech.* **2022**, *119*, 631–646. doi:10.1007/s00170-021-08235-4.
22. Kishore K, Chauhan SR, Sinha MK. A comprehensive investigation on eco-benign grindability improvement of Inconel 625 using nano-MQL. *Precis. Eng.* **2024**, *90*, 81–95. doi:10.1016/j.precisioneng.2024.08.004.
23. Li D, Duan Z, Hu X, Zhang D, Zhang Y. Automated classification and detection of multiple pavement distress images based on deep learning. *J. Traffic Transp. Eng.* **2023**, *10*, 276–290. doi:10.1016/j.jtte.2021.04.008.
24. Lu HZ, Li CF, Chen WM, Jiang ZJ. A single shot multibox detector based on welding operation method for biometrics recognition in smart cities. *Pattern Recogn. Lett.* **2020**, *140*, 295–302. doi:10.1016/j.patrec.2020.10.016.
25. Zhuang XL, Zhang TM. Detection of sick broilers by digital image processing and deep learning. *Biosyst. Eng.* **2019**, *179*, 106–116. doi:10.1016/j.biosystemseng.2019.01.003.
26. Amemiya S, Takao H, Kato S, Yamashita H, Sakamoto N, Abe O. Automatic detection of brain metastases on contrast-enhanced CT with deep-learning feature-fused single-shot detectors. *Eur. J. Radiol.* **2021**, *136*, 109577. doi:10.1016/j.ejrad.2021.109577.
27. Liu JM, Prabuwo AS, Abulfaraj AW, Miniaoui S, Taheri N. Cognitive cloud framework for waste dumping analysis using

- deep learning vision computing in healthy environment. *Comput. Electr. Eng.* **2023**, *110*, 108814. doi:10.1016/j.compeleceng.2023.108814.
28. Xie LF, Xiang X, Xu HN, Wang L, Lin LJ, Yin GF. FFCNN: A Deep Neural Network for Surface Defect Detection of Magnetic Tile. *IEEE T Ind. Electron.* **2021**, *68*, 3506–3516. doi:10.1109/Tie.2020.2982115.
29. Tabernik D, Sela S, Skvarc J, Skocaj D. Segmentation-based deep-learning approach for surface-defect detection. *J. Intell. Manuf.* **2020**, *31*, 759–776. doi:10.1007/s10845-019-01476-x.
30. Neuhauser FM, Bachmann G, Hora P. Surface defect classification and detection on extruded aluminum profiles using convolutional neural networks. *Int. J. Mater. Form.* **2020**, *13*, 591–603. doi:10.1007/s12289-019-01496-1.
31. Le HF, Zhang LJ, Liu YX. Surface Defect Detection of Industrial Parts Based on YOLOv5. *IEEE Access.* **2022**, *10*, 130784–130794. doi:10.1109/Access.2022.3228687.
32. Niu SL, Peng YR, Li B, Qiu YH, Niu TZ, Li WF. A novel deep learning motivated data augmentation system based on defect segmentation requirements. *J. Intell. Manuf.* **2024**, *35*, 687–701. doi:10.1007/s10845-022-02068-y.
33. Yu RY, Guo BY, Yang K. Selective Prototype Network for Few-Shot Metal Surface Defect Segmentation. *IEEE T Instrum. Meas.* **2022**, *71*, 3196447. doi:10.1109/Tim.2022.3196447.
34. Qian W, Zhu Z, Zhu C, Luo W, Zhu Y. Efficient deployment of Single Shot Multibox Detector network on FPGAs. *Integration* **2024**, *99*, 102255. doi:10.1016/j.vlsi.2024.102255.
35. Cai J, Makita Y, Zheng Y, Takahashi S, Hao W, Nakatoh Y. Single shot multibox detector for honeybee detection. *Comput. Electr. Eng.* **2022**, *104*, 108465. doi:10.1016/j.compeleceng.2022.108465.
36. Qiang J, Liu W, Li X, Guan P, Du Y, Liu B, et al. Detection of citrus pests in double backbone network based on single shot multibox detector. *Comput. Electron. Agric.* **2023**, *212*, 108158. doi:10.1016/j.compag.2023.108158.
37. Zhu W, Zhang H, Eastwood J, Qi X, Jia J, Cao Y. Concrete crack detection using lightweight attention feature fusion single shot multibox detector. *Knowl. -Based Syst.* **2023**, *261*, 110216. doi:10.1016/j.knosys.2022.110216.
38. Sun H, Xu H, Liu B, He D, He J, Zhang H, et al. MEAN-SSD: A novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks. *Comput. Electron. Agric.* **2021**, *189*, 106379. doi:10.1016/j.compag.2021.106379.
39. Shen YF, Zhou HL, Li JT, Jian FJ, Jayas DS. Detection of stored-grain insects using deep learning. *Comput. Electron. Agr.* **2018**, *145*, 319–325. doi:10.1016/j.compag.2017.11.039.
40. Erhan D, Szegedy C, Toshev A, Anguelov D. Scalable Object Detection Using Deep Neural Networks. IEEE. 2013. Available online: <https://arxiv.org/pdf/1312.2249> (accessed on 21 August 2024).
41. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.
42. Yu F, Koltun V, Funkhouser T. Dilated Residual Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017. doi:10.48550/arXiv.1705.09914.
43. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proc. Mach. Learn. Res.* **2015**, *37*, 448–456. doi:10.48550/arXiv.1502.03167.